

# Assisted Common Information: Further Results

Vinod M. Prabhakaran

École Polytechnique Fédérale de Lausanne  
Switzerland

Manoj M. Prabhakaran

University of Illinois, Urbana-Champaign  
Urbana, IL 61801

**Abstract**—We presented assisted common information as a generalization of Gács-Körner (GK) common information at ISIT 2010. The motivation for our formulation was to improve upperbounds on the efficiency of protocols for secure two-party sampling (which is a form of secure multi-party computation). Our upperbound was based on a monotonicity property of a rate-region (called the assisted residual information region) associated with the assisted common information formulation.

In this note we present further results. We explore the connection of assisted common information with the Gray-Wyner system. We show that the assisted residual information region and the Gray-Wyner region are connected by a simple relationship: the assisted residual information region is the increasing hull of the Gray-Wyner region under an affine map. Several known relationships between GK common information and Gray-Wyner system fall out as consequences of this. Quantities which arise in other source coding contexts acquire new interpretations.

In previous work we showed that assisted common information can be used to derive upperbounds on the rate at which a pair of parties can *securely sample* correlated random variables, given correlated random variables from another distribution. Here we present an example where the bound derived using assisted common information is much better than previously known bounds, and in fact is tight. This example considers correlated random variables defined in terms of standard variants of oblivious transfer, and is interesting on its own as it answers a natural question about these cryptographic primitives.

## I. INTRODUCTION

If  $U, V, W$  are independent random variables, a natural measure of “common information” of  $X = (U, V)$  and  $Y = (U, W)$  is  $H(U)$ . Observers of either  $X$  or  $Y$  may produce the common part  $U$  and conditioned on this common part, there is no residual information, i.e.,  $I(X; Y|U) = 0$ . Gács-Körner (GK) common information [5], [16] is a generalization of this to arbitrary  $X, Y$ . Two observers see  $X^n = (X_1, X_2, \dots, X_n)$  and  $Y^n = (Y_1, Y_2, \dots, Y_n)$ , resp., where  $(X_i, Y_i)$  are independent draws of  $(X, Y)$ . The observers produce  $W_1 = f_1(X^n)$  and  $W_2 = f_2(X^n)$  which have an asymptotically vanishing probability of not matching. GK common information is the largest entropy rate (normalized by  $n$ ) of such a common random variable. It was however shown that this value is the largest  $H(U)$  for which the random variables can be written as  $X = (U, V)$  and  $Y = (U, W)$  (where  $U, V, W$  may be dependent), i.e., the definition captures only an explicit form of common information in a single instance of  $X, Y$ .

At ISIT 2010 we presented a generalization of GK common information [13]. In our setup (see Figure 1), an omniscient genie (who has access to the  $X$  and  $Y$  sequences) assists

the users in generating the common random variables by sending them messages over rate-limited noiseless links. A three-dimensional trade-off region which characterizes the trade-off between the rates of the two noiseless links and the resulting residual information (defined as the conditional mutual information between the source sequences conditioned on the common random variable normalized by the length of the sequence) was derived. We call this the *assisted residual information region*. When the links have zero rates, we recover GK common information.

Our motivation for this generalization was an application to cryptography. Distributed dependent random variables are an important resource in the cryptographic task of secure multi-party computation. A fundamental problem here is for two parties to securely generate a certain pair of random variables, given another pair of random variables, by means of a protocol. Our main result there was that the assisted residual dependency region of the views of two parties engaged in such a protocol can only monotonically expand and not shrink which immediately leads to upperbounds on the efficiency with which a target pair of random variables can be generated from another pair. This work generalized previous work on monotones [17]. These works are in the same vein as [1], [4], [15], [10], [9], [7], [2], [14] which employ information theory to derive bounds on efficiency in cryptography.

In the first part of this paper we explore connections between the assisted common information system and the Gray-Wyner source coding system of [6]. In the Gray-Wyner system, a pair of sources is decomposed into three components: one public and two private. Using the public and one of the private components, one of the pair of sources must be recoverable, while the other source must be recoverable using the other private component and the public component. Gray-Wyner region is a three-dimensional region which characterizes the trade-offs between the rates at which the three components can be encoded.

We show that the assisted residual information region and the Gray-Wyner region are connected by a simple relationship: the assisted residual information region is the increasing hull<sup>1</sup> of the Gray-Wyner region under an affine map. Several known relationships between GK common information and Gray-Wyner system fall out as consequences of this. This also leads

<sup>1</sup>Increasing hull  $i(S)$  of a set  $S \subseteq \mathbb{R}^d$  is the set of all  $s \in \mathbb{R}^d$  such that there is a  $s' \in S$  such that  $s \geq s'$ , where the inequality is component-wise.

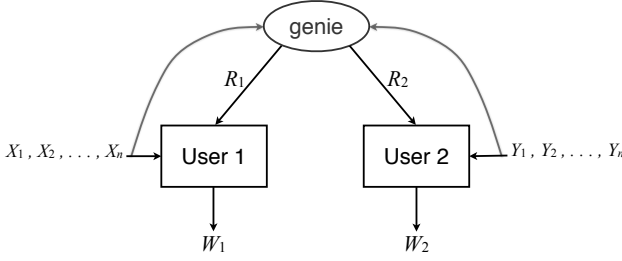


Fig. 1: Setup for assisted common information system. The users generate  $W_1$  and  $W_2$  which are required to agree with high probability. A genie assists the users by sending separate messages to them over rate-limited noiseless links. When the genie is absent the setup reduces to the one for Gács-Körner common information.

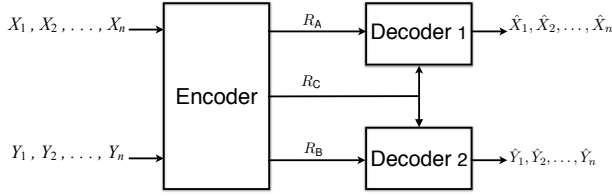


Fig. 2: Setup for Gray-Wyner (GW) system.

to alternative interpretations (in terms of the assisted common information system) to quantities which arise naturally in certain other source coding contexts. However, it must be noted that the Gray-Wyner region itself does not possess the monotonicity property which makes it less-suited for the cryptographic application which motivated [13].

The second half of the paper is a sequel to the cryptographic application in [13]. There we showed an example where our upperbound (on the efficiency with which a pair of random variables can be securely generated from another pair) strictly improved upon bounds from previous results. That example was contrived to highlight the shortcomings of prior work. Here we give yet another example where the upperbound from our result strictly improves on the prior work, but is further interesting for two reasons: firstly, the new example is based on natural correlated random variables that are widely studied (namely, variants of oblivious transfer), and secondly the new upperbound we can prove actually matches an easy lowerbound and is therefore tight.

## II. PRELIMINARIES

### A. Assisted Common Information System

We presented the following generalization of GK common information at ISIT, 2010 [13]. We call it the assisted common information system.

Consider Figure 1. For a pair of random variables  $(X, Y)$ , we say that a rate pair  $(R_1, R_2)$  *enables* a residual information rate  $R_{RD}$  if for every  $\epsilon > 0$ , there is a large enough integer  $n$  and (deterministic) functions  $f_k : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \{1, \dots, 2^{n(R_k + \epsilon)}\}$ ,  $(k = 1, 2)$ ,  $g_1 : \mathcal{X}^n \times \{1, \dots, 2^{n(R_1 + \epsilon)}\} \rightarrow$

$\mathbb{Z}$ , and  $g_2 : \mathcal{Y}^n \times \{1, \dots, 2^{n(R_2 + \epsilon)}\} \rightarrow \mathbb{Z}$  (where  $\mathbb{Z}$  is the set of integers) such that

$$\Pr(g_1(X^n, f_1(X^n, Y^n)) \neq g_2(Y^n, f_2(X^n, Y^n))) \leq \epsilon, \quad (1)$$

$$\frac{1}{n} I(X^n; Y^n | g_1(X^n, f_1(X^n, Y^n))) \leq R_{RD} + \epsilon. \quad (2)$$

*Definition 2.1:* We define the *assisted residual information region*<sup>2</sup>  $\mathcal{R}_{ACI}(X, Y)$  of a pair of random variables  $(X, Y)$  with joint distribution  $p_{X,Y}$  as the set of all  $(r_1, r_2, r_{RD}) \in \mathbb{R}_+^3$  for which there is a  $(R_1, R_2, R_{RD})$  such that  $r_1 \geq R_1$ ,  $r_2 \geq R_2$ ,  $r_{RD} \geq R_{RD}$ , and  $(R_1, R_2)$  enables the residual information rate  $R_{RD}$ . In other words,

$$\mathcal{R}_{ACI}(X, Y) \triangleq i(\{(R_1, R_2, R_{RD}) : (R_1, R_2) \text{ enables } R_{RD}\}),$$

where  $i(S)$  denotes the *increasing hull* of  $S \subseteq \mathbb{R}_+^3$ :  $i(S) = \{s \in \mathbb{R}_+^3 : s \geq s' \text{ component-wise for some } s' \in S\}$ .

We will write  $\mathcal{R}_{ACI}$  when the random variables involved are obvious from the context.

When the two rates from the genie are zero, we recover Gács-Körner common information,  $C_{GK}$  [5], [16]. Let  $R_{RD-0} \triangleq \inf\{R_{RD} : (0, 0, R_{RD}) \in \mathcal{R}_{ACI}(X, Y)\}$ . Then we have the following proposition.

*Proposition 2.1:*

$$C_{GK}(X, Y) = I(X; Y) - R_{RD-0}. \quad (3)$$

Further

$$R_{RD-0} = \inf_{p_{U|XY} : I(X; U|Y) = I(Y; U|X) = 0} I(X; Y|U) \quad (4)$$

which gives

$$C_{GK}(X, Y) = \sup_{p_{U|XY} : I(X; U|Y) = I(Y; U|X) = 0} H(U). \quad (5)$$

Moreover,  $C_{GK}(X, Y) = 0$  unless there are  $X', Y', U'$  such that  $X = (X', U')$ ,  $Y = (Y', U')$ , in which case  $C_{GK} = \max_{U' : X=(X', U'), Y=(Y', U')} H(U')$ .

The proof of this proposition and all other results are available in the appendix. The proof of (4) relies on the following characterization of  $\mathcal{R}_{ACI}$  which was proved in [13]. Let  $\mathcal{P}$  be the set of all marginal p.m.f's  $p_{U|X,Y}$  such that the cardinality of alphabet  $\mathcal{U}$  of  $U$  is  $|\mathcal{X}||\mathcal{Y}| + 2$ .

*Proposition 2.2:*

$$\mathcal{R}_{ACI}(X, Y) = i\left(\bigcup_{p_{U|X,Y} \in \mathcal{P}} \{(I(Y; U|X), I(X; U|Y), I(X; Y|U))\}\right)$$

<sup>2</sup>We may also define an analogous *assisted common information region* by replacing the definition in (2) by

$$\frac{1}{n} I(X^n, Y^n; g_1(X^n, f_1(X^n, Y^n))) \geq R_{CI} - \epsilon.$$

See [13] for this and its connection to the above definition. In effect, the definitions are equivalent as we discuss there. We work with assisted residual information region since it has a simple monotonicity property (Theorem 4.1) which makes it appealing for deriving bounds for secure two-party sampling.

### B. Gray-Wyner system

The Gray-Wyner system is shown in Figure 2. It is a source coding problem formulated as follows: We say that a rate 3-tuple  $(R_A, R_B, R_C)$  is *achievable* if for every  $\epsilon > 0$ , there is a large enough integer  $n$  and (deterministic) encoder functions  $f_A : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \{1, \dots, 2^{n(R_A+\epsilon)}\}$ ,  $f_B : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \{1, \dots, 2^{n(R_B+\epsilon)}\}$ ,  $f_C : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \{1, \dots, 2^{n(R_C+\epsilon)}\}$ , and (deterministic) decoder functions  $g_{AC} : \{1, \dots, 2^{n(R_A+\epsilon)}\} \times \{1, \dots, 2^{n(R_C+\epsilon)}\} \rightarrow \mathcal{X}^n$ , and  $g_{BC} : \{1, \dots, 2^{n(R_B+\epsilon)}\} \times \{1, \dots, 2^{n(R_C+\epsilon)}\} \rightarrow \mathcal{Y}^n$  such that

$$\Pr(g_{AC}(f_A(X^n, Y^n), f_C(X^n, Y^n)) \neq X^n) \leq \epsilon, \quad (6)$$

$$\Pr(g_{BC}(f_B(X^n, Y^n), f_C(X^n, Y^n)) \neq Y^n) \leq \epsilon. \quad (7)$$

**Definition 2.2:** The Gray-Wyner region  $\mathcal{R}_{\text{GW}}(X, Y)$  is the set of all achievable rate 3-tuples.

We write  $\mathcal{R}_{\text{GW}}$  when the random variables are clear from the context. A simple lower-bound to  $\mathcal{R}_{\text{GW}}(X, Y)$  is

$$\mathcal{L}_{\text{GW}}(X, Y) = \{(R_A, R_B, R_C) : R_A + R_C \geq H(X), R_B + R_C \geq H(Y), R_A + R_B + R_C \geq H(X, Y)\} \quad (8)$$

The Gray-Wyner region was characterized in [6].

**Proposition 2.3 ([6]):**

$$\mathcal{R}_{\text{GW}}(X, Y) = i \left( \bigcup_{p_{U|X,Y} \in \mathcal{P}} \{(H(X|U), H(Y|U), I(X, Y; U))\} \right)$$

The Gray-Wyner system generalizes the setup for Wyner's common information [19] which is defined as the smallest  $R_C$  such that the outputs of the encoder taken together is an asymptotically efficient representation of  $(X, Y)$ , i.e., when  $R_A + R_B + R_C = H(X, Y)$ . Using the above proposition we have

**Proposition 2.4:**

$$\begin{aligned} C_{\text{Wyner}}(X, Y) &= \inf\{R_C : (R_A, R_B, R_C) \in \mathcal{R}_{\text{GW}}(X, Y), \\ &\quad R_A + R_B + R_C = H(X, Y)\} \\ &= \inf_{p_{U|X,Y} \in \mathcal{P}: X-U-Y} I(X, Y; U) \end{aligned}$$

### C. Known connections

The following connections between the two systems are known:

- Gács-Körner common information can be obtained from the Gray-Wyner region [3, Problem 4.28, pg. 404].

$$C_{\text{GK}}(X, Y) = \sup\{R_C : R_A + R_C = H(X), R_B + R_C = H(Y), (R_A, R_B, R_C) \in \mathcal{R}_{\text{GW}}\} \quad (9)$$

Alternatively [11],

$$C_{\text{GK}}(X, Y) = \sup\{R : R \leq I(X; Y), \{R_C = R\} \cap \mathcal{L}_{\text{GW}} \subseteq \mathcal{R}_{\text{GW}}\} \quad (10)$$

- Wyner's common information can be obtained from the Gács-Körner system [13, Corollary 2.3].

$$C_{\text{Wyner}}(X, Y) = I(X; Y) + \inf_{(R_1, R_2, 0) \in \mathcal{R}_{\text{ACI}}} R_1 + R_2. \quad (11)$$

### III. RELATIONSHIP BETWEEN ASSISTED COMMON INFORMATION AND GRAY-WYNER SYSTEMS

**Theorem 3.1:** Let  $\mathcal{R}'_{\text{GW}}(X, Y)$  be the image of  $\mathcal{R}_{\text{GW}}(X, Y)$  under the affine map  $f_{X,Y}$  defined below.

$$f_{X,Y} \left( \begin{bmatrix} R_A \\ R_B \\ R_C \end{bmatrix} \right) \triangleq \begin{bmatrix} R_A + R_C - H(X) \\ R_B + R_C - H(Y) \\ R_A + R_B + R_C - H(X, Y) \end{bmatrix}.$$

Then

$$\mathcal{R}_{\text{ACI}}(X, Y) = i(\mathcal{R}'_{\text{GW}}(X, Y)).$$

Thus, the assisted residual information region  $\mathcal{R}_{\text{ACI}}(X, Y)$  is the increasing hull of the Gray-Wyner region  $\mathcal{R}_{\text{GW}}(X, Y)$  under an affine map  $f_{X,Y}$ . The map, in fact, computes the gap of  $\mathcal{R}_{\text{GW}}(X, Y)$  to the simple lower bound  $\mathcal{L}_{\text{GW}}(X, Y)$  of (8) under a coordinate transformation. The first coordinate of  $\mathcal{R}'_{\text{GW}}$  is indeed the gap between the (sum) rate at which the first decoder in the Gray-Wyner system receives data and the minimum possible rate at which it may receive data so that it can losslessly reproduce  $X^n$ . The second coordinate has a similar interpretation with respect to the second decoder. The third coordinate is the gap between the rate at which the encoder sends data and the minimum possible rate at which it may transmit to allow both decoders to losslessly reproduce their respective sources.

It must, however, be noted that the Gray-Wyner region itself does not possess the monotonicity property of  $\mathcal{R}_{\text{ACI}}$  which leads to Theorem 4.1 and is therefore less-suited for the cryptographic application which motivated [13].

The two points we noted in Section II-C fall out of Theorem 3.1.

**Corollary 3.2:**

$$\begin{aligned} C_{\text{GK}}(X, Y) &= \sup\{R_C : R_A + R_C = H(X), R_B + R_C = H(Y), (R_A, R_B, R_C) \in \mathcal{R}_{\text{GW}}(X, Y)\} \quad (9) \\ &= \sup\{R : R \leq I(X; Y), \{R_C = R\} \cap \mathcal{L}_{\text{GW}}(X, Y) \subseteq \mathcal{R}_{\text{GW}}(X, Y)\} \quad (10) \end{aligned}$$

**Corollary 3.3:**

$$C_{\text{Wyner}}(X, Y) = I(X; Y) + \inf_{(R_1, R_2, 0) \in \mathcal{R}_{\text{ACI}}(X, Y)} R_1 + R_2. \quad (11)$$

Analogous to the definition of  $R_{\text{RD-0}}$ , we define the axes intercepts on the other two axes.

$$\begin{aligned} R_{1-0} &\triangleq \inf\{R_1 : (R_1, 0, 0) \in \mathcal{R}_{\text{ACI}}\} \\ R_{2-0} &\triangleq \inf\{R_2 : (0, R_2, 0) \in \mathcal{R}_{\text{ACI}}\} \end{aligned}$$

$R_{1-0}$  (resp.,  $R_{2-0}$ ) is the rate at which the genie must communicate when it has a link to only the user who receives  $X$  (resp.  $Y$ ) source so that the users can produce a common random variable conditioned on which the sources are independent<sup>3</sup>. Using Proposition 2.2 we can show that

$$R_{1-0} = \inf_{p_{U|X,Y} \in \mathcal{P}: I(X;U|Y)=I(X;Y|U)=0} I(Y;U|X), \quad (12)$$

$$R_{2-0} = \inf_{p_{U|X,Y} \in \mathcal{P}: I(Y;U|X)=I(X;Y|U)=0} I(X;U|Y). \quad (13)$$

These quantities were identified in [17] and shown to possess a monotonic property in the context of secure two-party sampling (a result which [13] generalized).

As we will show below, this pair of quantities is closely related to a pair which has been identified elsewhere in the context of lossless coding with side-information [12] and the Gray-Wyner system [11]. Let (following the notation of [11])

$$\begin{aligned} G(Y \rightarrow X) &= \inf\{R_C : (H(X|Y), H(Y) - R_C, R_C) \in \mathcal{R}_{\text{GW}}(X, Y)\}, \\ G(X \rightarrow Y) &= \inf\{R_C : (H(X) - R_C, H(Y|X), R_C) \in \mathcal{R}_{\text{GW}}(X, Y)\}. \end{aligned}$$

It has been shown [12], [11] that  $G(Y \rightarrow X)$  is the smallest rate at which side-information  $Y$  may be coded and sent to a decoder which is interested in recovering  $X$  with asymptotically vanishing probability of error if the decoder receives  $X$  coded and sent at a rate of only  $H(X|Y)$  (which is the minimum possible rate which will allow such recovery). Further, [11] arrives at the maximum of  $G(Y \rightarrow X)$  and  $G(X \rightarrow Y)$  as a dual to the alternative definition of  $C_{\text{GK}}$  in (10) from the Gray-Wyner system.

We have the following relationship between the two pairs of quantities.

*Corollary 3.4:*

$$G(Y \rightarrow X) = I(X; Y) + R_{1-0}, \quad (14)$$

$$G(X \rightarrow Y) = I(X; Y) + R_{2-0}. \quad (15)$$

Further,

$$\inf\{R : R \geq I(X; Y), (R_C = R) \cap \mathcal{L}_{\text{GW}}(X, Y) \subseteq \mathcal{R}_{\text{GW}}(X, Y)\} \text{repeated Minkowski sum} \\ = \max(G(Y \rightarrow X), G(X \rightarrow Y)) \quad (16)$$

$$= I(X; Y) + \max(R_{1-0}, R_{2-0}). \quad (17)$$

#### IV. CRYPTOGRAPHIC APPLICATION

The cryptographic problem we consider is of 2-party *secure sampling*: Alice and Bob should sample correlated random variables  $(U, V)$  (Alice getting  $U$  and Bob getting  $V$ ), such that Alice's view during the sampling protocol reveals nothing more to her about Bob's outcome  $V$  than what her own outcome  $U$  reveals to her, and similarly Bob's view reveals nothing more about Alice's outcome than is revealed by his

own outcome. This is an important special case of *secure multi-party computation*, a central problem in modern cryptography.

However, it is well-known (see for instance [18] and references therein) that very few distributions can be sampled from in this way, unless the computation is aided by a *set up* — some correlated random variables that are given to the parties at the beginning of the protocol. The set up itself will be from some distribution  $(X, Y)$  (Alice gets  $X$  and Bob gets  $Y$ ) which is different from the desired distribution  $(U, V)$ . The fundamental question then is, which set ups  $(X, Y)$  can be used to securely sample which distributions  $(U, V)$ , and *how efficiently*.

We restrict ourselves to the setting of *honest-but-curious* players. In this case, the requirements on a protocol  $\Pi$  for securely sampling  $(U, V)$  given a set up  $(X, Y)$  can be stated as follows, in terms of the outputs and the views of the parties from the protocol:<sup>4</sup>

$$\begin{aligned} (\Pi_{\text{Alice}}^{\text{out}}(X, Y), \Pi_{\text{Bob}}^{\text{out}}(X, Y)) &= (U, V) \\ \Pi_{\text{Alice}}^{\text{view}}(X, Y) &\leftrightarrow \Pi_{\text{Alice}}^{\text{out}}(X, Y) \leftrightarrow \Pi_{\text{Bob}}^{\text{out}}(X, Y) \\ \Pi_{\text{Alice}}^{\text{out}}(X, Y) &\leftrightarrow \Pi_{\text{Bob}}^{\text{out}}(X, Y) \leftrightarrow \Pi_{\text{Bob}}^{\text{view}}(X, Y) \end{aligned}$$

These three conditions correspond to correctness, security against a curious Alice and security against a curious Bob, respectively.

In [13], we showed that the region  $\mathcal{R}_{\text{ACI}}$  can be used as a measure of cryptographic complexity of correlated random variables (a smaller region  $\mathcal{R}_{\text{ACI}}$  corresponding to a higher complexity), in that the rate at which a pair  $(U, V)$  can be securely sampled given a set up  $(X, Y)$  can be upperbounded by the ratio of their complexity measures. More formally, there we presented the following result. (For completeness, a proof is provided in the appendix.)

*Theorem 4.1 ([13]):* If  $n_1$  independent copies of a pair of correlated random variables  $(U, V)$  can be securely realized from  $n_2$  independent copies of a pair of correlated random variables  $(X, Y)$ , then  $n_1 \mathcal{R}_{\text{ACI}}(X, Y) \subseteq n_2 \mathcal{R}_{\text{ACI}}(U, V)$  (where multiplication by  $n$  refers to  $n$ -times repeated Minkowski sum).

In [13] we gave an instance of pairs  $(U, V)$  and  $(X, Y)$  such that the upperbound on the rate at which instances of  $(U, V)$  can be securely sampled from instances of  $(X, Y)$  that is implied by the above result strictly improved on the upperbounds that could be derived from previous results. These pairs were contrived to highlight the shortcomings of prior work. Here we give yet another example where the upperbound from our result strictly improves on prior work, but is further interesting for two reasons: firstly, the new example is based on natural correlated random variables that are widely studied (namely, variants of oblivious transfer), and secondly, the new upperbound we can prove actually matches an easy lowerbound and is therefore tight.

<sup>3</sup>Though the definition allows for zero-rate communication to the other user and a zero-rate (but non-zero) residual conditional mutual information, it can be shown from the expression for these rates in (12)-(13) that there is a scheme which achieves exact conditional independence and requires no communication to the other user.

<sup>4</sup>Here we state the conditions for “perfect security,” but our definitions and results generalize to the setting of “statistical security,” where a small statistical error is allowed.

## A. A New Example

We now discuss the new example where our upperbound is not only strictly better than the previously best available upperbound, but is also tight.

*Example 4.1:* Let  $S_{A,1}, S_{A,2}, S_{B,1}, S_{B,2} \in \{0,1\}^L$  and  $C_A, C_B \in \{1,2\}$  be six independent random variables all of which are uniformly distributed over their alphabets. Consider a pair of random variables  $X, Y$  defined as  $X = (C_A, S_{A,1}, S_{A,2}, S_{B,C_A})$  and  $Y = (C_B, S_{B,1}, S_{B,2}, S_{A,C_B})$ . Notice that these are in fact a pair of independent string-oblivious transfers (string-OT's) of string length  $L$  in opposite directions. Let  $U, V$  be a pair of random variables whose joint distribution is the same as that of  $X, Y$ , but with  $L = 1$ . In other words,  $U, V$  are a pair of independent bit-OT's in opposite directions. The goal is to characterize the efficiency with which we may securely generate independent instances of  $U, V$  from independent instances of  $X, Y$  for  $L > 1$ . Here efficiency is the supremum of  $n_2/n_1$  over secure sampling schemes which produce  $n_2$  independent copies of  $(U, V)$  from  $n_1$  independent copies of  $(X, Y)$ .

It is easy to see that  $\mathcal{R}_{ACI}(X, Y)$  intersects the co-ordinate axes at  $(1+L, 0, 0)$ ,  $(0, 1+L, 0)$ , and  $(0, 0, 2L)$ . From, these we can immediately obtain the upperbound of [17] on the efficiency, namely  $(1+L)/2$ . Notice that this is dependent on  $L$  and would suggest that (several) long string-OT pairs can be turned into several (more) bit-OT pairs. However, as we show below, the efficiency of conversion is just 1, i.e., the best one can do is to turn each pair of string-OT's into a pair of bit-OT's.

We will show that  $\inf\{R_1 + R_2 : (R_1, R_2, 0) \in \mathcal{R}_{ACI}(U, V)\} = 2$ . But,  $(1, 1, 0) \in \mathcal{R}_{ACI}(X, Y)$ . This can be seen by setting  $Q = (C_A, C_B, S_{A,C_B}, S_{B,C_A})$  for which  $(R_1, R_2, R_{RD}) = (1, 1, 0)$ . Thus,  $\inf\{R_1 + R_2 : (R_1, R_2, 0) \in \mathcal{R}_{ACI}(X, Y)\} \leq 2$ . Hence, from Theorem 4.1, we may conclude that the efficiency of conversion we are after is 1.

It only remains to characterize  $\inf\{R_1 + R_2 : (R_1, R_2, 0) \in \mathcal{R}_{ACI}(U, V)\}$ . The following lemma, which is proved in the appendix, provides the required characterization.

*Lemma 4.2:*

$$\inf\{R_1 + R_2 : (R_1, R_2, 0) \in \mathcal{R}_{ACI}(U, V)\} = 2.$$

## ACKNOWLEDGEMENTS

The authors would like to gratefully acknowledge discussions with Venkat Anantharam, Péter Gács, and Young-Han Kim. The example in Section IV-A is based on a suggestion by Jürg Wullschleger.

## REFERENCES

- [1] D. Beaver, "Correlated pseudorandomness and the complexity of private computations," in *Proc. 28th STOC*, pp. 479–488. ACM, 1996.
- [2] I. Csizsár and R. Ahlswede, "On oblivious transfer capacity," in *Proc. International Symposium on Information Theory (ISIT)*, pp. 2061–2064, 2007.
- [3] I. Csizsár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Akadémiai Kiadó, Budapest, 1981.
- [4] Y. Dodis and S. Micali, "Lower bounds for oblivious transfer reductions," in Jacques Stern, editor, *EUROCRYPT*, vol. 1592 of *Lecture Notes in Computer Science*, pp. 42–55. Springer, 1999.

- [5] P. Gács and J. Körner, "Common information is far less than mutual information," *Problems of Control and Information Theory*, 2(2):119–162, 1973.
- [6] R. M. Gray and A. D. Wyner, "Source coding for a simple network," *Bell System Technical Journal*, vol. 53, pp. 1681–1721, 1974.
- [7] H. Imai, K. Morozov, and A. C. A. Nascimento, "On the oblivious transfer capacity of the erasure channel," in *Proc. International Symposium on Information Theory (ISIT)*, pp. 1428–1431, 2006.
- [8] H. Imai, K. Morozov, and A. C. A. Nascimento, "Efficient oblivious transfer protocols achieving a non-zero rate from any non-trivial noisy correlation," in *International Conference on Information Theoretic Security (ICITS)*, 2007.
- [9] H. Imai, K. Morozov, A. C. A. Nascimento, and A. Winter, "Efficient protocols achieving the commitment capacity of noisy correlations," in *International Symposium on Information Theory (ISIT)*, pp. 1432–1436, 2006.
- [10] H. Imai, J. Müller-Quade, A. C. A. Nascimento, and A. Winter, "Rates for bit commitment and coin tossing from noisy correlation," in *International Symposium on Information Theory (ISIT)*, pp. 45–, 2004.
- [11] S. Kamath and V. Anantharam, "A new dual to the Gács-Körner common information defined via the Gray-Wyner system," in *Proc. 48th Allerton Conf. on Communication, Control, and Computing*, pp. 1340–1346, 2010.
- [12] D. Marco and M. Effros, "On lossless coding with coded side information," *IEEE Transactions on Information Theory*, vol. 55, no. 7, pp. 3284–3296, 2009.
- [13] V. M. Prabhakaran and M. Prabhakaran, "Assisted common information," in *International Symposium on Information Theory (ISIT)*, pp. 2602–2606, 2010. Extended draft available at <http://arxiv.org/abs/1002.1916>.
- [14] S. Winkler and J. Wullschleger, "Statistical impossibility results for oblivious transfer reductions," Cryptology ePrint Archive, Report 2009/508, 2009. <http://eprint.iacr.org/>.
- [15] A. Winter, A. C. A. Nascimento, and H. Imai, "Commitment capacity of discrete memoryless channels," In Kenneth G. Paterson, editor, *IMA Int. Conf.*, vol. 2898 of *Lecture Notes in Computer Science*, pp. 35–51. Springer, 2003.
- [16] H. S. Witsenhausen, "On sequences of pairs of dependent random variables," *SIAM Journal of Applied Mathematics*, 28:100–113, 1975.
- [17] S. Wolf and J. Wullschleger, "New monotones and lower bounds in unconditional two-party computation," *IEEE Transactions on Information Theory*, 54(6):2792–2797, 2008.
- [18] J. Wullschleger, Oblivious-Transfer Amplification. Ph.D. thesis, Swiss Federal Institute of Technology, Zürich. <http://arxiv.org/abs/cs/0608076>.
- [19] A. D. Wyner, "The common information of two dependent random variables," *IEEE Transactions on Information Theory*, 21(2), 163–179, 1975.

## APPENDIX

### *Proof of Proposition 2.1:*

GK common information  $C_{GK}$  is defined as the supremum of the set of  $R$  such that for every  $\epsilon > 0$  there are maps  $g_1 : \mathcal{X}^n \rightarrow \mathbb{Z}$ , and  $g_2 : \mathcal{Y}^n \rightarrow \mathbb{Z}$  for a sufficiently large  $n$  which satisfy

$$\Pr(g_1(X^n) \neq g_2(Y^n)) \leq \epsilon, \quad (18)$$

$$\frac{1}{n} H(g_1(X^n)) \geq R - \epsilon. \quad (19)$$

An alternative definition which allows for a genie with zero-rate links to the users is given below. It is easy to see that this can only lead to a larger value. But as we will show, the definitions are in fact equivalent.

Let  $C'_{GK}$  be the supremum of the set of  $R$  such that for every  $\epsilon > 0$  there are maps  $f_k : \mathcal{X}^n \times \mathcal{Y}^n \rightarrow \{1, \dots, 2^{n\epsilon}\}$ , ( $k = 1, 2$ ),  $g_1 : \mathcal{X}^n \times \{1, \dots, 2^{n\epsilon}\} \rightarrow \mathbb{Z}$ , and  $g_2 : \mathcal{Y}^n \times \{1, \dots, 2^{n\epsilon}\} \rightarrow \mathbb{Z}$  for a sufficiently large  $n$  which satisfy (1) and

$$\frac{1}{n} H(g_1(X^n, f_1(X^n, Y^n))) \geq R - \epsilon.$$

$$I(Y; U|X) = I(X, Y; U) - I(X; U) = H(X|U) + I(X, Y; U) - H(X), \quad (20)$$

$$I(X; U|Y) = I(X, Y; U) - I(Y; U) = H(Y|U) + I(X, Y; U) - H(Y), \text{ and} \quad (21)$$

$$I(X; Y|U) = H(X|U) + H(Y|U) - H(X, Y|U) = H(X|U) + H(Y|U) + I(X, Y; U) - H(X, Y). \quad (22)$$

Clearly,  $C'_{\text{GK}} \geq C_{\text{GK}}$ . We first show

$$C'_{\text{GK}} = I(X; Y) - R_{\text{RD-0}}. \quad (20)$$

Let  $U = g_1(X^n(f_1(X^n, Y^n)))$ . Then

$$\begin{aligned} I(X^n; Y^n|U) + H(U) &= I(X^n; Y^n|U) + I(X^n, Y^n; U) \\ &= H(X^n, Y^n) - H(X^n|Y^n, U) - H(Y^n|X^n, U) \\ &= I(X^n; Y^n) + I(X^n; U|Y^n) + I(Y^n; U|X^n) \\ &\geq nI(X; Y). \end{aligned}$$

Therefore, if the maps satisfy (2), then

$$\begin{aligned} H(U) &\geq nI(X; Y) - I(X^n; Y^n|U) \\ &\geq nI(X; Y) - n(R_{\text{RD}} + \epsilon) \\ &= n(I(X; Y) - R_{\text{RD}} - \epsilon) \end{aligned}$$

which implies (20).

With  $C_{\text{GK}}$  replaced by  $C'_{\text{GK}}$ , we can prove (4)-(5) as follows: (4) follows from Proposition 2.2; (4) and (3) imply (5). See [13, section II.B] for a proof from (5) of the explicit characterization stated at the end of the proposition. Since this explicit form can be achieved without any communication from the genie, it follows that  $C'_{\text{GK}} = C_{\text{GK}}$ . ■

*Proof of Theorem 3.1:*

It is easy to prove the above theorem from the single-letter expressions for the regions in propositions 2.2 and 2.3 by making use of the mutual information equalities (20)-(22) at the top of the page. ■

*Proof of Corollary 3.2:*

$$\begin{aligned} \sup\{R_C : R_A + R_C = H(X), \\ R_B + R_C = H(Y), (R_A, R_B, R_C) \in \mathcal{R}_{\text{GW}}\} \\ \stackrel{(a)}{=} \sup\{R : (0, 0, I(X; Y) - R) \in \mathcal{R}'_{\text{GW}}\} \\ \stackrel{(b)}{=} \sup\{R : (0, 0, I(X; Y) - R) \in \mathcal{R}_{\text{ACI}}\}, \end{aligned}$$

where (a) follows from the definition  $\mathcal{R}'_{\text{GW}} = f(\mathcal{R}_{\text{GW}})$ . The  $\leq$  direction of (b) follows directly from Theorem 3.1. But  $<$  cannot hold since if  $(0, 0, I(X; Y) - R) \in \mathcal{R}_{\text{ACI}}$ , then there is a  $R' \geq R$  such that  $(0, 0, I(X; Y) - R') \in \mathcal{R}'_{\text{GW}}$ . Finally, (c) follows from Proposition 2.1.

To arrive at the alternative form, we verify the equivalence of the two forms.

$$\begin{aligned} \{R : R \leq I(X; Y), \{R_C = R\} \cap \mathcal{L}_{\text{GW}} \subseteq \mathcal{R}_{\text{GW}}\} \\ = \{R_C : R_A + R_C = H(X), \\ R_B + R_C = H(Y), (R_A, R_B, R_C) \in \mathcal{R}_{\text{GW}}\}. \end{aligned}$$

$\subseteq$ : if  $R \leq I(X; Y)$ , then  $(H(X) - R, H(Y) - R, R) \in \{R_C = R\} \cap \mathcal{L}_{\text{GW}}$ .

$\supseteq$ : Let  $s = (H(X) - R_C, H(Y) - R_C, R_C) \in \mathcal{R}_{\text{GW}}$ . Then (a)  $R_C \leq I(X; Y)$  since  $s \in \mathcal{L}_{\text{GW}}$ , and (b) if  $s' = (r_A, r_B, R_C) \in \mathcal{L}_{\text{GW}}$ , then since  $r_A \geq H(X) - R_C$  and  $r_B \geq H(Y) - R_C$ , we have  $s' \geq s$  (component-wise) which implies that  $s' \in \mathcal{R}_{\text{GW}}$  from the definition of the **GW** system. ■

*Proof of Corollary 3.3:*

$$\begin{aligned} C_{\text{Wyner}} &= \inf\{R_C : (R_A, R_B, R_C) \in \mathcal{R}_{\text{GW}}, \\ &\quad R_A + R_B + R_C = H(X, Y)\} \\ &\stackrel{(a)}{=} \inf\{R_1 + R_2 + I(X; Y) : (R_1, R_2, 0) \in \mathcal{R}'_{\text{GW}}\} \\ &\stackrel{(b)}{=} \inf\{R_1 + R_2 + I(X; Y) : (R_1, R_2, 0) \in \mathcal{R}_{\text{ACI}}\}, \end{aligned}$$

where (a) follows from the definition  $\mathcal{R}'_{\text{GW}} = f(\mathcal{R}_{\text{GW}})$ ; (b) follows from Theorem 3.1:  $\geq$  direction follows directly from the theorem. But  $>$  cannot hold, since by the theorem, if  $(R_1, R_2, 0) \in \mathcal{R}_{\text{ACI}}$  then there exists  $(R'_1, R'_2, 0) \in \mathcal{R}'_{\text{GW}}$  such that  $R'_1 \leq R_1$  and  $R'_2 \leq R_2$ . ■

*Proof of Corollary 3.4:*

$$\begin{aligned} G(Y \rightarrow X) &= \inf\{R_C : (H(X|Y), H(Y) - R_C, R_C) \in \mathcal{R}_{\text{GW}}\}, \\ &\stackrel{(a)}{=} \inf\{R : (R - I(X; Y), 0, 0) \in \mathcal{R}'_{\text{GW}}\} \\ &\stackrel{(b)}{=} \inf\{R : (R - I(X; Y), 0, 0) \in \mathcal{R}_{\text{ACI}}\} \\ &\stackrel{(c)}{=} I(X; Y) + R_{1-0}, \end{aligned}$$

where (a) follows from  $\mathcal{R}'_{\text{GW}} = f(\mathcal{R}_{\text{GW}})$ . (b) is a consequence of Theorem 3.1: And (c) follows from the definition of  $R_{1-0}$ .

Similarly we get (15). The equality (16) is proved in [11] which along with (14)-(15) implies (17). ■

*Proof of Theorem 4.1:* The theorem is in fact corollary 3.2 of [13] which follows immediately from Theorem 3.1 of [13] and the following lemma:

*Lemma A.1:* Let the pair of random variables  $(X_1, Y_1)$  be independent of the pair  $(X_2, Y_2)$ . If  $X = (X_1, X_2)$  and  $Y = (Y_1, Y_2)$ , then

$$\mathcal{R}_{\text{ACI}}(X, Y) = \mathcal{R}_{\text{ACI}}(X_1, Y_1) + \mathcal{R}_{\text{ACI}}(X_2, Y_2).$$

For completeness, we give a proof of Theorem 3.1 of [13] below since the proof was not provided there. This also contains a proof of Lemma A.1 (see (d) below). Please refer [13] for notation and a statement of the theorem being proved below.

We show that under each step of a secure protocol,  $\mathcal{R}_{\text{ACI}}$  can only grow. ■

(a) *Local computation cannot shrink it:* For all random variables with  $X - Y - Z$ , we have  $\mathcal{R}_{\text{ACI}}(X, YZ) \supseteq \mathcal{R}_{\text{ACI}}(X, Y)$  and  $\mathcal{R}_{\text{ACI}}(XY, Z) \supseteq \mathcal{R}_{\text{ACI}}(X, Y)$ .

The first set inclusion follows from the fact that for the joint p.m.f.  $p_{X,Y,Z,Q} = p_{X,Y}p_{Z|Y}p_{Q|X,Y}$

$$I(X; Y, Z|Q) = I(X; Y|Q)$$

$$I(Q; Y, Z|X) = I(Q; Y|X)$$

$$I(X; Q|Y, Z) = I(X; Q|Y).$$

(b) *Communication cannot shrink it:* For all random variables  $(X, Y)$  and functions  $f$  over the support of  $X$  (resp.  $Y$ ), we have  $\mathcal{R}_{\text{ACI}}(X, (Y, f(X))) \supseteq \mathcal{R}_{\text{ACI}}(X, Y)$  (resp.  $\mathcal{R}_{\text{ACI}}((X, f(Y)), Y) \supseteq \mathcal{R}_{\text{ACI}}(X, Y)$ ).

The first set inclusion follows from the following facts for the joint p.m.f.  $p_{X,Y,Z,Q} = p_{X,Y}p_{Z|Y}p_{Q|X,Y}$ :

$$\begin{aligned} I(X; Y, f(X)|Q, f(X)) &= I(X; Y|Q, f(X)) \\ &\leq I(X; Y|Q) \end{aligned}$$

$$\begin{aligned} I(X; Q, f(X)|Y, f(X)) &= I(X; Q|Y, f(X)) \\ &\leq I(X; Q|Y) \end{aligned}$$

$$I(Y; Q, f(X)|X) = I(Y; Q|X)$$

(c) *Securely derived outputs do not have a smaller region:* For all random variables  $X, U, V, Y$  such that  $X - U - V$  and  $U - V - Y$ , we have  $\mathcal{R}_{\text{ACI}}(U, V) \supseteq \mathcal{R}_{\text{ACI}}((X, U), (Y, V))$ .

This follows from the following facts for (dependent) random variables  $X, Y, U, V, Q$  which satisfy the Markov chains  $X - U - V$  and  $U - V - Y$ :

$$\begin{aligned} I(X, U; Y, V|Q) &\geq I(U; V|Q), \\ I(X, U; Q|Y, V) &= I(X, U; Q, Y|V) - I(X, U; Y|V) \\ &\stackrel{(a)}{=} (I(U; Q, Y|V) + I(X; Q, Y|U, V)) \\ &\quad - I(X; Y|U, V) \\ &\geq I(U; Q|V), \end{aligned}$$

and similarly

$$I(Y, V; Q|X, U) \geq I(V; Q|U),$$

where we used  $U - V - Y$  to obtain equality (a).

(d) *Regions of independent pairs add up:* If  $(X, Y)$  is independent of  $(U, V)$ , we have  $\mathcal{R}_{\text{ACI}}((X, U), (Y, V)) = \mathcal{R}_{\text{ACI}}(X, Y) + \mathcal{R}_{\text{ACI}}(U, V)$ . This follows easily from the following facts:

For the joint p.m.f.  $p_{X,Y}p_{U,V}p_{Q_1|X,Y}p_{Q_2|U,V}$ , we have

$$I(X, U; Y, V|Q_1, Q_2) = I(X; Y|Q_1) + I(U, V|Q_2)$$

$$I(X, U; Q_1, Q_2|Y, V) = I(X; Q_1|Y) + I(U; Q_2|V)$$

$$I(Y, V; Q_1, Q_2|X, U) = I(Y; Q_1|X) + I(V; Q_2|U)$$

And, for the joint p.m.f.  $p_{X,Y}p_{U,V}p_{Q|X,Y,U,V}$ , we have

$$I(X, U; Y, V|Q) \geq I(X; Y|Q) + I(U; V|Q)$$

$$I(X, U; Q|Y, V) \geq I(X; Q|Y) + I(U; Q|V)$$

$$I(Y, V; Q|X, U) \geq I(Y; Q|X) + I(V; Q|U)$$

*Proof of Lemma 4.2:*

By Lemma A.1, we need only characterize the  $\inf\{R_1 + R_2 : (R_1, R_2, 0) \in \mathcal{R}_{\text{ACI}}\}$  of one of the pair of independent bit-OT's. Let us denote one bit-OT by  $A, B$ : where  $A = (S_1, S_2) \in \{0, 1\}^2$  uniformly distributed over its alphabet and  $B = (C, S_C)$ , where  $C \in \{1, 2\}$  is independent of  $A$  and uniformly distributed over its alphabet. By Proposition 2.2,

$$\begin{aligned} &\inf\{R_1 + R_2 : (R_1, R_2, 0) \in \mathcal{R}_{\text{ACI}}(A, B)\} \\ &= \inf_{p_{Q|A,B} \in \mathcal{P}: I(A; B|Q)=0} I(B; Q|A) + I(A; Q|B) \\ &= H(A|B) + H(B|A) \\ &\quad - \sup_{p_{Q|A,B} \in \mathcal{P}: I(A; B|Q)=0} H(A|Q, B) + H(B|Q, A). \end{aligned}$$

We show below that the sup term is 1. Since  $H(A|B) + H(B|A) = 2$ , this will allow us to conclude that the smallest sum-rate of  $\mathcal{R}_{\text{RD}}(0)$  of  $A, B$  is 1. Invoking the lemma above, the corresponding smallest sum-rate for  $U, V$  is then 2 as required.

To show that the sup term is 1, notice that the only valid choices of  $p_{Q|A,B}$  are such that  $I(A; B|Q) = 0$ . This means that the resulting  $p_{A,B|Q}(\cdot, \cdot|q)$  must belong to one of eight possible classes shown in Figure 3b (for any  $q$  with non-zero probability  $p_Q(q)$ ; we may assume that all  $q$ 's have non-zero probability without loss of generality). Recall that there is a cardinality bound on  $Q$ ; let us denote the alphabet of  $Q$  by  $\{q_1, q_2, \dots, q_N\}$ , where  $N$  is the cardinality bound.

We will first show that there is no loss of generality in assuming that no more than one of the  $q_i$ 's is such that its  $p_{A,B|Q}(\cdot, \cdot|q_i)$  belongs to the same class (and hence we may take  $N = 8$ ). Suppose,  $q_1$  and  $q_2$  belong to the same class, say class 1, with parameters  $p_1$  and  $p_2$  respectively. Then, if we denote the binary entropy function by  $H_2(\cdot)$ , we have

$$\begin{aligned} &H(A|Q, B) + H(B|Q, A) \\ &= \sum_{k=1}^N p_Q(q_k) (H(A|B, Q = q_k) + H(B|A, Q = q_k)) \\ &= p_Q(q_1)H_2(p_1) + p_Q(q_2)H_2(p_2) \\ &\quad + \sum_{k=3}^N p_Q(q_k) (H(A|B, Q = q_k) + H(B|A, Q = q_k)) \\ &\leq (p_Q(q_1) + p_Q(q_2)) H_2\left(\frac{p_Q(q_1)p_1 + p_Q(q_2)p_2}{p_Q(q_1) + p_Q(q_2)}\right) \\ &\quad + \sum_{k=3}^N p_Q(q_k) (H(A|B, Q = q_k) + H(B|A, Q = q_k)), \end{aligned}$$

where the inequality (Jensen's) follows from the concavity of the binary entropy function. Thus, we can define a  $Q'$  of alphabet size  $N - 1$  where letters  $q_1, q_2$  are replaced by  $q_0$  such that  $p_{Q'}(q_0) = p_Q(q_1) + p_Q(q_2)$ , and  $p_{A,B|Q'=q_0}$  is in class 1 with parameter  $\frac{p_Q(q_1)p_1 + p_Q(q_2)p_2}{p_Q(q_1) + p_Q(q_2)}$ , while maintaining for  $i = 3, \dots, N$ ,  $p_{Q'}(q_i) = p_Q(q_i)$  and  $p_{A,B|Q'}(a, b|q_i) = p_{A,B|Q}(a, b|q_i)$ . (It is easy to verify (a) that this gives a valid

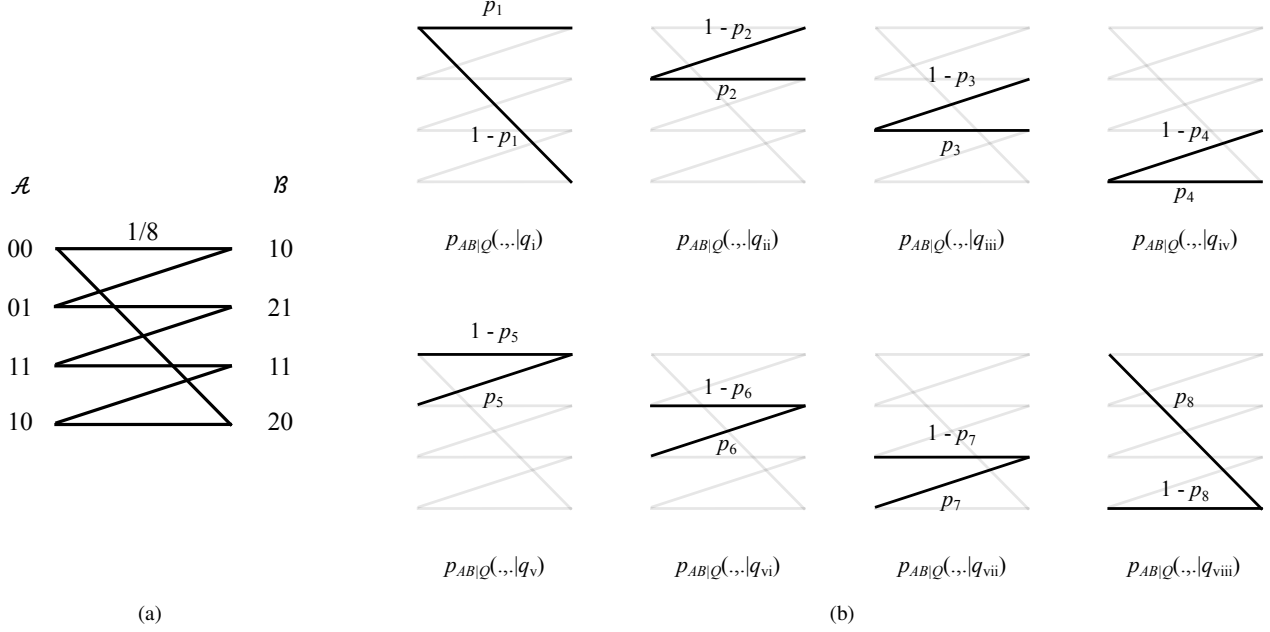


Fig. 3: (a) Joint p.m.f. of  $A, B$ . Each solid line represents a probability mass of  $1/8$ . (b) Eight possible classes that  $p_{A,B|Q}(\cdot, \cdot | q)$  may belong to for a  $p_{Q|A,B}$  which satisfies  $I(A; B|Q) = 0$ .

joint p.m.f. for  $p_{A,B,Q'}$ , (b) that the induced  $p_{A,B}$  is the same as the original, and (c) that the induced  $p_{Q'|A,B}$  satisfies the condition  $I(A; B|Q') = 0$ .) Then, the above inequality states that

$$H(A|Q, B) + H(B|Q, A) \leq H(A, Q', B) + H(B|Q', A)$$

proving our claim.

Thus, without loss of generality, we may assume that  $N = 8$  and  $p_{A,B|Q}(\cdot, \cdot | q_i)$  belongs to class  $i$ . Notice that

$$\begin{aligned} p_{Q|A,B}(q_1|00, 10) + p_{Q|A,B}(q_5|00, 10) &= 1, \\ p_{Q|A,B}(q_2|01, 10) + p_{Q|A,B}(q_5|01, 10) &= 1, \\ p_{Q|A,B}(q_2|01, 21) + p_{Q|A,B}(q_6|01, 21) &= 1, \\ p_{Q|A,B}(q_3|11, 21) + p_{Q|A,B}(q_6|11, 21) &= 1, \\ p_{Q|A,B}(q_3, 11, 11) + p_{Q|A,B}(q_7|11, 11) &= 1, \\ p_{Q|A,B}(q_4|10, 11) + p_{Q|A,B}(q_7|10, 11) &= 1, \\ p_{Q|A,B}(q_4|10, 20) + p_{Q|A,B}(q_8|10, 20) &= 1, \\ p_{Q|A,B}(q_1|00, 20) + p_{Q|A,B}(q_8|00, 20) &= 1. \end{aligned}$$

Let us define

$$\begin{aligned} \tilde{p}_1 &\triangleq p_{Q|A,B}(q_1|00, 10), & \tilde{p}_5 &\triangleq p_{Q|A,B}(q_5|01, 10), \\ \tilde{p}_2 &\triangleq p_{Q|A,B}(q_2|01, 21), & \tilde{p}_6 &\triangleq p_{Q|A,B}(q_6|11, 21), \\ \tilde{p}_3 &\triangleq p_{Q|A,B}(q_3|11, 11), & \tilde{p}_7 &\triangleq p_{Q|A,B}(q_7|10, 11), \\ \tilde{p}_4 &\triangleq p_{Q|A,B}(q_4|10, 20), & \tilde{p}_8 &\triangleq p_{Q|A,B}(q_8|00, 20). \end{aligned}$$

Let us evaluate  $H(B|Q, A)$  in terms of the above parameters.

Notice that  $H(B|Q = q_i, A) = 0$  for  $i = 5, \dots, 8$ . Hence

$$\begin{aligned} H(B|Q, A) &= \sum_{(q,a) \in \{(1,00), (2,01), (3,11), (4,10)\}} p_{Q,A}(q, a) H(B|Q = q, A = a) \\ &= \frac{\tilde{p}_1 + (1 - \tilde{p}_8)}{8} H_2\left(\frac{\tilde{p}_1}{\tilde{p}_1 + (1 - \tilde{p}_8)}\right) \\ &\quad + \frac{\tilde{p}_2 + (1 - \tilde{p}_5)}{8} H_2\left(\frac{\tilde{p}_2}{\tilde{p}_2 + (1 - \tilde{p}_5)}\right) \\ &\quad + \frac{\tilde{p}_3 + (1 - \tilde{p}_6)}{8} H_2\left(\frac{\tilde{p}_3}{\tilde{p}_3 + (1 - \tilde{p}_6)}\right) \\ &\quad + \frac{\tilde{p}_4 + (1 - \tilde{p}_7)}{8} H_2\left(\frac{\tilde{p}_4}{\tilde{p}_4 + (1 - \tilde{p}_7)}\right) \\ &\leq \frac{4 + \sum_{i=1}^4 \tilde{p}_i - \sum_{j=5}^8 \tilde{p}_j}{8}, \end{aligned}$$

where the inequality follows from the fact that binary entropy function is upperbounded by 1. Similarly, we can get

$$H(A|Q, B) \leq \frac{4 + \sum_{j=5}^8 \tilde{p}_j - \sum_{i=1}^4 \tilde{p}_i}{8}.$$

Combining, we obtain the desired

$$H(B|Q, A) + H(A|Q, B) \leq 1.$$

■